COMPONENT PART NOTICE

THIS PAPER IS A COMPONENT PART OF THE FOLLOWING COMPILATION REPORT:

TITLE: The Practical Assessment of Pilot Workload: Flight Mechanics Panel of AGARD.

TO ORDER THE COMPLETE COMPILATION REPORT, USE AD-A184 834.

THE COMPONENT PART IS PROVIDED HERE TO ALLOW USERS ACCESS TO INDIVIDUALLY AUTHORED SECTIONS OF PROCEEDING, ANNALS, SYMPOSIA, ETC. HOWEVER, THE COMPONENT SHOULD BE CONSIDERED WITHIN THE CONTEXT OF THE OVERALL COMPILATION REPORT AND NOT AS A STAND-ALONE TECHNICAL REPORT.

THE FOLLOWING COMPONENT PART NUMBERS COMPRISE THE COMPILATION REPORT:

AD#: P005 629 thru P005 643	AD#:
AD#:	AD#:
AD#:	AD#:

Accession For	
NTIS GRA&I	☒
DTIC TAB	☐
Unannounced	☐
Justification	
By	
Distribution/	
Availability Codes	
Dist	Avail and/or Special
A-1	

AD-P005 629

CHAPTER 2

IN-FLIGHT WORKLOAD ASSESSMENT USING EMBEDDED SECONDARY RADIO COMMUNICATIONS TASKS

by

Clark A Shingledecker
Ergometrics Technology, Inc
4401 Dayton-Xenia Road
Dayton Ohio USA

INTRODUCTION

Traditional Secondary Task Measures

A widely accepted conceptual framework which forms the basis for many workload measurement techniques represents the human operator as a limited capacity information processing system. According to this general model, workload may be defined as the degree to which the operator's processing capacity is occupied by mental activities. Overload, and resulting performance decrement, occurs when capacity is insufficient to meet task demands. Since the momentary capacity of the operator is unknown and submaximal workload levels cannot be inferred from his or her performance on the task of interest, an indirect measure can be obtained by evaluating the amount of spare capacity available under a given set of task conditions.

The behavioral approach to assessing spare capacity involves the use of the secondary task technique. In this method, operators are given an additional information processing task to perform in conjunction with the task of interest. The rationale underlying the use of secondary tasks is that by applying an extra load which produces a total information processing demand that exceeds the operator's capacity, workload can be measured by observing the difference between single task and dual task performances. As noted by Ogden, Levine, and Eisner (1), secondary tasks can be employed in two ways. Used as a loading techniques, the method requires subjects to perform the secondary task under all circumstances with the intent of displaying overload effects in primary task performance. When secondary tasks are used as a workload measure, performance on the primary task is emphasized and secondary task performance is observed as an index of the workload of the primary task. Although specific research questions may require a choice of one of these applications, combined task decrement may also be used as an estimate of mutual interference and workload (2).

Unlike time-based analytical methods, the secondary task approach to assessing spare mental capacity has the potential for being sensitive to the degree of mental effort or attention devoted to information processing as well as to the temporal aspects of workload. The secondary task technique has the further advantage of producing a measure based on task performance, which is the variable that all workload measures ultimately must predict if they are to be of any value.

Although secondary task methodology has proven to be a useful technique for the investigation of cognitive processes, its practical application as a workload measurement tool has often been confined to the earliest stages of aircraft system design. As Schiflett (3) has noted, most workload measures have been developed for, and are most applicable to, the laboratory environment in which highly controlled, part task studies of workload can be conducted. When subsystems are combined to evaluate mission performance in the context of high fidelity simulations or flight tests, many workload assessment methods become difficult to employ because they are impractical or present potential safety hazards. As a result, workload measurement at the critical later stages of system development is often performed using relatively informal and qualitative techniques.

Three specific problems are encountered when traditional laboratory secondary tasks are considered for use during advanced development of aircraft. One practical problem is the physical instrumentation of the secondary task. In a flight test environment, and to a lesser extent in a simulator, introducing or adding any extra equipment to the crew station may be unacceptable. Even when sufficient space can be reserved, the possibility of obstruction or distraction caused by the additional instrumentation can limit the feasibility of using the secondary task.

A second problem with the implementation of secondary tasks is the possibility of intrusion on primary flight duties. Although some performance decrement may be tolerable, task interference can easily complicate the interpretation of data in test environments where measures of all performance variables may be unavailable. A more serious consequence of primary task intrusion in the flight test environment is the potential for compromising flight safety.

The final factor limiting the use of secondary task measures is operator acceptance (4). Whether used to induce stress or to measure reserve capacity, a secondary task is likely to produce misleading data if the operator fails to integrate it with his normal duties. Acceptance is a potential problem with all laboratory tasks because they are obvious, artificial additions to the crewstation and have little face validity or congruence with the general performance situation. Such test conditions can lead the operator to neglect the secondary task or, because of its novelty, allow it to assume an artificially high priority. Thus, lack of operator acceptance can become a major contributor to primary task intrusion as well as a source of measurement error.

Embedded Secondary Tasks

The embedded secondary task methodology was developed by Shingledecker et al (5) (6) to improve the practical utility of dual task measures for in-flight workload assessment, while retaining many of the scientific advantages associated with traditional laboratory secondary tasks. The concept of the embedded secondary task is based on the hypothesis that instrumentation limitations, task intrusion, and poor operator acceptance can be minimized by designing secondary tasks which are fully integrated with system hardware and with the crewmember's conception of the mission environment. By their nature, such tasks are realistic components of crewstation activity, yet their performance can be manipulated and measured independently of the primary activities of interest.

While several classes of aircrew activity are potential candidates for isolation and use as embedded tasks, radio communications tasks are particularly suitable for this purpose. The radio communications which are most useful as embedded tasks are those initiated by a message sent from another aircraft or a ground controller to a pilot whose workload is to be assessed. Upon detection and identification of a relevant message, the pilot must engage in a sequence of verbal responses and radio switching activities in order to meet the demands of the communicated request.

Such tasks closely resemble the nonadaptive discrete secondary tasks used in numerous workload studies and have many properties of good measurement tasks. Communications call upon a wide variety of information processing abilities and can be varied along several dimensions of complexity. Furthermore, no auxiliary crewstation equipment is necessary to control the experiment or to collect performance data. The opportunity for obstruction or peripheral interference is also minimized since the auditory channel is not shared by other tasks and verbal responses are generally unique to radio communications activities, while switch actions can be dealt with by the pilot's free hand. Most importantly, communications tasks are an integral part of a pilot's in-flight duties. As a result, lengthy training requirements are eliminated and high face validity is achieved. Additionally, the realistic nature of the activity makes artificial task interactions improbable because the pilot has predetermined priorities assigned to communications and other cockpit functions. These features make communications activities especially valuable for use as secondary tasks since pilots consider them to be important, but will normally devote less attention to communications as more crucial tasks become difficult to perform.

DESCRIPTION OF THE TECHNIQUE

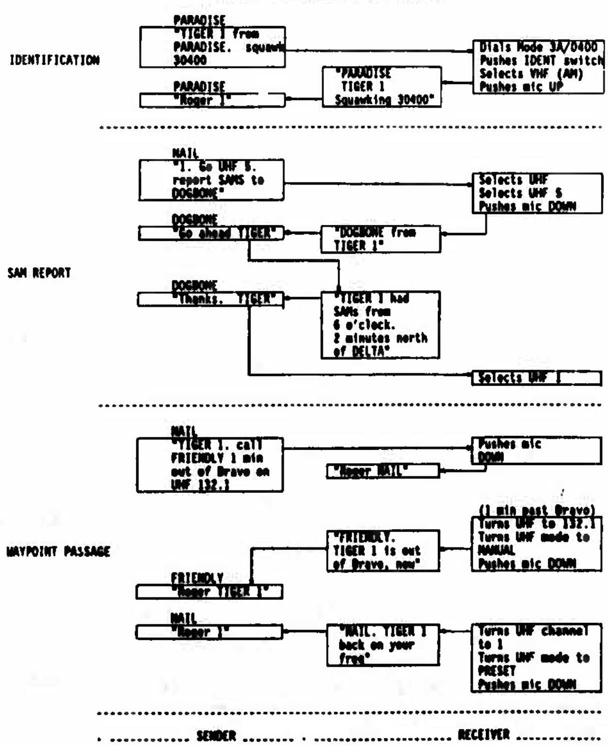
Task Selection

The use of radio communications activities as embedded workload measures for high fidelity simulation or in-flight environments requires careful selection of the communications tasks to insure both realism and valid measurement. First, a group of candidate tasks must be identified which are relevant to the aircraft and mission of interest. Appropriate tasks may be obtained by interviewing operational pilots. In documenting these tasks, particular care should be taken to specify all verbiage used by the sender and receiver of the radio messages as well as the manual control actions required of the aircraft member. Additionally, the typical frequency and time of occurrence for each task should be noted.

Tasks which do not appear in the majority of interview responses or which vary in procedure among protocols should be eliminated from the group. Furthermore, those tasks which tend to take precedence over normal aircraft control functions should be avoided. For example, messages communicating threat would undoubtedly alter a pilot's normal attentional priorities and would shift any workload induced performance decrement to primary flight tasks. Some sample tasks which were obtained from single set fighter attack pilot and which meet the requirements discussed above are shown in Figure 1.

FIGURE 1.

SAMPLE SECONDARY COMMUNICATIONS TASKS FROM AN AIR TO GROUND ATTACK SCENARIO



Traditional discrete trial laboratory secondary tasks insure comparability of individual data points by repeating identical stimuli. Since communications tasks are not obviously comparable in their information processing demands, a second step that must be taken in task selection is workload scaling. Such scaling permits the experimenter to select a realistic combination of tasks for use in workload measurement which present equivalent estimated subsidiary loading levels. Shingledecker and his co-workers (4) evaluated three alternative apriori scaling techniques to achieve this purpose. Of the analytical and subjective methods which were tested, an information theoretical approach produced the highest correlation with dual task decrement scores.

This scaling technique is based on the assumption that the mental workload of communications tasks can be predicted by assessing the uncertainty associated with the reception of stimuli and execution of responses required of the pilot. Once a radio message is detected, the pilot must make two perceptual decisions to identify the intended receiver of the message and its sender. According to information theory, the demands associated with each decision can be estimated by determining the number of potential receivers and senders in the scenario and calculating a bit measure of the uncertainty of the decisions ($\log_2 N$). Thus, a message beginning with "Dogbone, this is Pounder. . . ." would require the reception of 2.32 bits if there were five active receivers on the radio channel, plus one bit if there were two active message senders in the scenario.

Following these perceptual decisions, the pilot must make action decisions in response to the instructions received. Action decisions may require verbal and/or manual responses, and again may be quantified by determining the number of alternative actions that could be made. Thus, if a UHF radio channel change were required, the action sequence might involve the selection of a tuning mode with two alternatives (1 bit), turning a rotary control to one of twenty preset channels (4.32 bits) and pressing a microphone switch with two-positions to acknowledge the message (1 bit).

While verbal response decisions are more difficult to quantify in the information theoretic metric, a majority of these behaviors can be classified into one of two types. The simplest activity is a message confirmation which involves simple information conservation. Within this scaling method such responses are assigned a value of one bit. The second type of response requires the pilot to select a new receiver from among those active in the scenario and to report some information from cockpit displays or the external visual scene. In these cases the verbal response requirements are computed by summing the bits associated with selecting from among the available receivers, and adding a single bit for the report.

An overall estimate of the loading presented by a communications task is derived by summing the bit values calculated for all perceptual decisions and for each manual and verbal action decision in the task sequence. While this quasi-information theoretic method relies on assumptions of equiprobability of alternatives and independence of sequential actions, empirical tests indicate that it provides a reasonable estimate of secondary communications task loading. Values calculated for a set of candidate tasks may be used to select tasks with approximately equal load for workload assessment within a single flight scenario.

Workload Assessment

Once usable communications tasks are identified, their application for workload measurement closely follows the procedure normally used for traditional secondary tasks. Prior to testing the aircrew subjects should be briefed on the workload assessment procedure, emphasizing that their responses to some of the communications messages that will occur during the flight will be used to measure workload. They should be told to respond to these messages in a normal fashion, and to maintain primary flight task performance under all conditions (ie, the communcations should not receive extra effort not afforded them in typical flying situations). Thus, they should respond to communications as quickly and accurately as possible, but not at the expense of primary flight control and management.

Prior to the test flight each participating pilot should review the communications tasks to be used for workload assessment. Finally, baseline single task performance should be recorded for each pilot on each of the tasks. This can be accomplished by presenting the tasks prior to take-off while the pilots are seated in the cockpits and are able to devote their full attention to the tasks. Performance scoring in both the single task baseline trials and in the in-flight test condition is accomplished by measuring each communication task completion time to the nearest 0.5 second. Times may be recorded manually beginning with the onset of the sender's message and ending with the final word of the pilot's response which completes the task sequence.

During the test flights the communications tasks should be presented to the pilots in accordance with a specified protocol developed to address the workload question of interest. Relative differences in workload between mission segments, cockpit design options etc are determined by comparing the magnitude of the difference between total task completion times for the baseline single task tests and the in-flight tests.

EXAMPLE OF USE

As in most other available workload measurement methods, the secondary communications task technique provides data which are interpreted in terms of comparisons among baseline conditions and various test conditions. Thus, no single example can address the potential range of workload questions or experimental design to which the technique is applicable. The example outlined below involves a hypothetical cockpit/system design issue. Equivalent examples could be developed to examine other comparative topics such as the impact on workload of flight experience, stressors or environmental conditions.

In the following case, the goal of the operational study will be to determine whether a new flight control system proposed for a twin jet transport aircraft reduces pilot workload during instrument approach and landing. It is assumed that previous test flights have revealed no objective evidence of major differences in flight performance between the current system and the proposed system. Two aircraft are available for the test, one equipped with the current flight control system and the other with the new system. Furthermore, five pilots who have equal flight time in the current system and have been thoroughly trained with the new system are available as test subjects.

Three types of communications tasks have been selected and scaled for use in the workload assessment. Each of these is initiated by air traffic control, but could be presented by an on board observer whose microphone is patched-in to the radios. The three messages are: 1) a request for radio frequency change (eg "FLYWAY 219, Contact approach on 118.1"), 2) a request to change transponder codes (eg "FLYWAY 219, Squawk 5133", 3) a request for traffic information (eg "FLYWAY 219, do you see DELTA 1011?").

The pilots are briefed on appropriate response procedures and single task baseline performance is timed before the test flights. Each pilot flies the standard approach and landing twice in the current aircraft and twice in the aircraft equipped with the new flight control system. The four flights are accomplished in a randomized order determined for each pilot. Data from any approach and landing which does not meet the flight performance requirements specified in the experimental protocol are rejected and the trial is repeated.

The secondary communication tasks are relayed to the pilot according to a predetermined schedule starting with the initial transition to approach and ending with the touchdown. Six tasks (two of each type) are presented in addition to normal communications during the final five minutes of flight. Performance is scored by computing the time difference between baseline single task performance for each communication task and the performance during each occurrence of the task in flight. Mean decrement scores are computed for each task under the current and proposed flight control system and proposed flight control system conditions. A statistically significant reduction in decrement scores when using the new system would be interpreted as evidence for improved workload as a result of the design change.

LIMITATIONS

Like other operational test methods, the embedded secondary communications task technique can present problems of experimental control and precision of measurement which may affect the sensitivity of a workload assessment. Consequently, its value as a realistic methodology should not be allowed to outweigh the need for preliminary testing under part task simulation conditions. Both laboratory measurements and confirmatory flight tests are required to provide a complete and defensible workload analysis. Specific issues that should be considered when deciding to employ this method for flight test purposes include:

1 At present, no standardized secondary communication tasks are available for general use. Each application requires selection and scaling of tasks which are tailored to individual workload questions, specific systems and their missions.

2 The technique produces relatively few data points per unit time. Each task requires several seconds to perform and must occur with a relatively realistic frequency. As a result, embedded communication tasks are more suited to evaluating workload over extended periods of five or more minutes than to brief intervals of interest.

3 The method has not been tested to determine the degree to which different tasks produce diagnostic measures of workload. That is, it is not known which communications tasks are most sensitive to particular types of crew station loading. Available data indicate that communications tasks requiring manual activities (eg, radio tuning) tend to provide optimal measures of crew workload in tasks which involve aircraft control as a primary component.

REFERENCES

1	OGDEN G D LEVINE J M EISNER E J	Measurement of workload by secondary tasks. Human Factors, 21, 529-548 1979
2	WICKENS C D MOUNTFORD S J SCHREINER W	Multiple resources, task-hemispheric integrity and individual differences in time-sharing. Human Factors, 23, 211-229 1981
3	SCHIFLETT S G	An Annotated Bibliography, US Naval Air Test Center, SY-257R-76 1976
4	SHINGLEDECKER C A et al	Subsidiary radio communications task for workload assessment in R&D simulations: I. Task development and workload scaling. Air Force Aerospace Medical Research Laboratory Technical Report, AFAMRL-TR-80-126 1980
5	SHINGLEDECKER C A CRABTREE M S	Subsidiary radio communications tasks for workload assessment in R&D simulations: II Task sensitivity evaluation. Air Force Aerospace Medical Research Laboratory Technical Report, AFAMRL-TR-82-57 1982